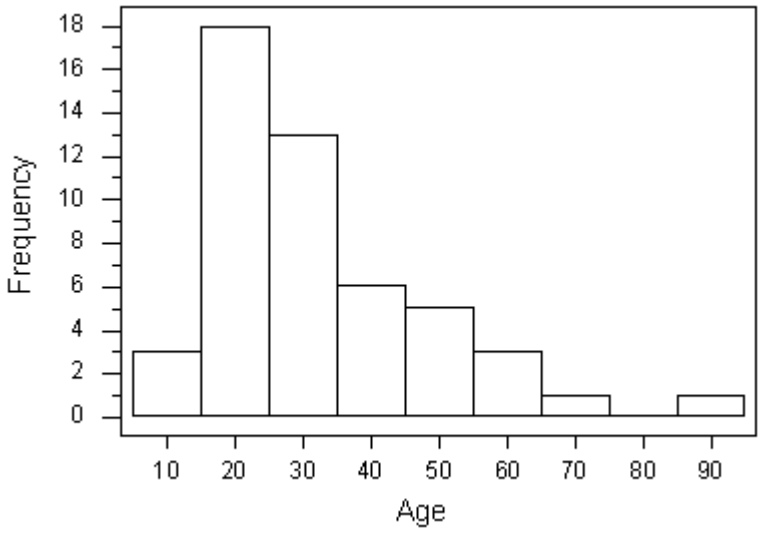
Examining Distributions Checkpoint 1

Pool1

The following 5 questions relate to the same histogram, shown below.

Question (1)

The histogram below displays the distribution of 50 ages at death due to trauma (accidents and homicides) that were observed in a certain hospital during a week.



What percentage of deaths were individuals younger than 35?

A: 34%
B: 60%
C: 68%
D: 70%
E: 80%

Feedback

A : 0

✗ This is not quite right. It's true that 3 + 18 + 13 = 34 out of the 50 fall below age 35 (those within the first three bars), however, note that the problem asks for *percentage* rather than count. (C) is the correct answer.

B : 0

✗ This is not quite right. While it's clear from looking at the histogram that more than 50% of the observations fall below 35 (i.e., within the first three bars), the percentage is actually higher than 60%. Count the number of observations below 35 and then use that number to calculate the percentage. (C) is the correct answer.

C : 10

✓ Good Job. From the histogram we can ï¬nd that 3 + 18 + 13 = 34 out of the 50 observations fall below age 35 (those within the first three bars), and since the question asks about the *percentage* of observations, the answer is (34/50) * 100 = 68%.

D : 0

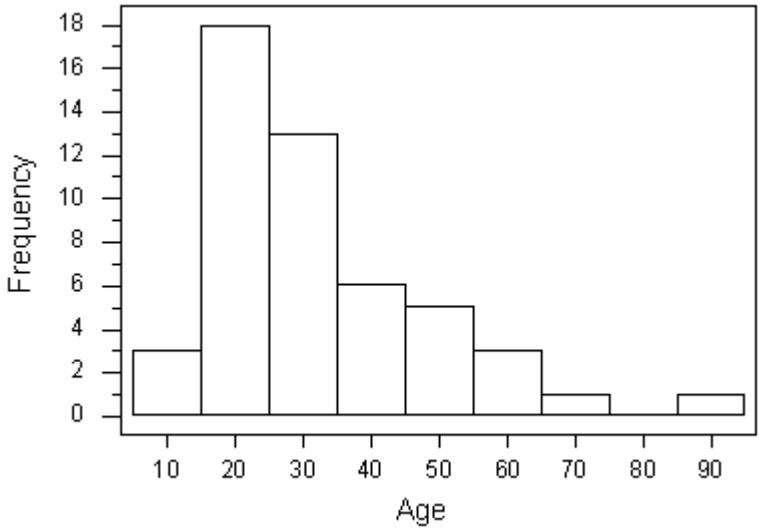
✗ This is not quite right. While it's clear from looking at the histogram that more than 50% of the observations fall below 35 (those within the first three bars), the percentage is actually lower than 70%. Count the number of observations below 35 and then use that number to calculate the percentage. (C) is the correct answer.

E : 0

✗ This is not quite right. While it's clear from looking at the histogram that more than 50% of the observations fall below 35 (those within the first three bars), this is not quite as high as 80%. Count the number of observations below 35 and then use that number to calculate the percentage. (C) is the correct answer.

Question (2)

Here again is the histogram showing the distribution of 50 ages at death due to trauma (accidents and homicides) that were observed in a certain hospital during a week.



Which of the following best describes the shape of the histogram?

A: Symmetric
B: Left-skewed with no outliers
C: Right-skewed with no outliers
D: Left-skewed with a possible outlier
E: Right-skewed with a possible outlier

Feedback

A : 0

✗ This is not quite right. The shape of the distribution is clearly skewed. Where is the "tail," and are there any possible outliers? (E) is the correct answer.

B : 0

✗ This is not quite right. While it is true that the distribution is skewed, recall that the direction of the skewness is determined by where the "tail" is. Also, look at the distribution carefully. Are there any possible outliers? (E) is the correct answer.

C : 0

✗ This is not quite right. Good job for recognizing that the distribution is skewed right, however note that one of the other answers more accurately describes the distribution. (E) is the correct answer.

D : 0

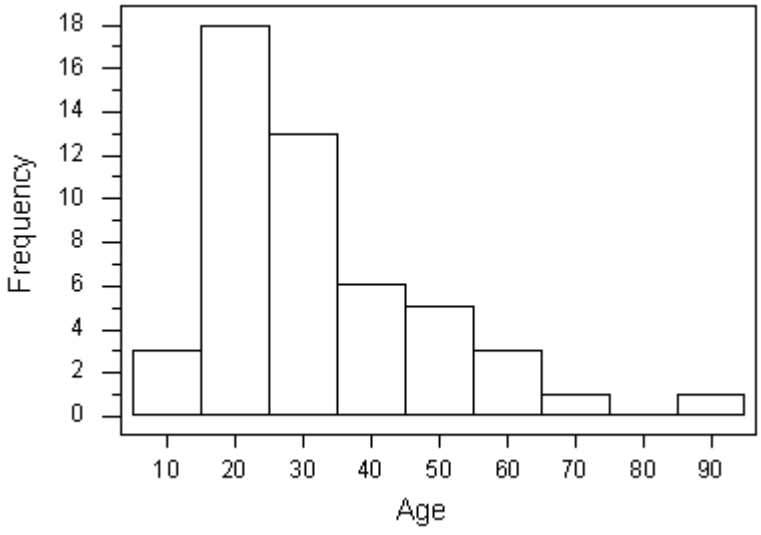
✗ This is not quite right. While it is true that the distribution is skewed, recall that the direction of the skewness is determined by where the "tail" is. (E) is the correct answer.

E : 10

✓ Good Job! Indeed, the distribution of ages is skewed right (which is where the "tail" of the distribution is). Also the largest observation—approximately equal to 90—is a possible high outlier.

Question (3)

Here again is the histogram showing the distribution of 50 ages at death due to trauma (accidents and homicides) that were observed in a certain hospital during a week.



For the data described by the above histogram,

A: the median will be bigger than the mean.
B: the median will be smaller than the mean.
C: the median and the mean will be about the same.
D: the median and the range will be about the same.
E: both (B) and (D) are correct.

Feedback

A : 0

✗ This is not quite right. Notice this distribution is skewed to the right (the tail is towards the right) and we also see a suspected outlier on the right side. Remember that both of these shape features will tend to "pull" the ***mean*** more than they would pull the median. (B) is the correct answer.

B : 10

✓ Good job! The ***mean*** is more sensitive than the median to the "tail" of a distribution, and also to outliers. We see that this distribution is skewed to the right (the tail is towards the right) and we also see a suspected outlier on the right side. Both of these shape features will tend to "pull" the ***mean*** more than they would pull the median.

C : 0

✗ This is not quite right. Remember that we'd expect the mean and the median to be nearly the same only if the distribution was nearly symmetric. But notice this distribution is skewed to the right (the tail is towards the right) and we also see a suspected outlier on the right side. Remember that both of these shape features will tend to "pull" the ***mean*** more than they would pull the median. (B) is the correct answer.

D : 0

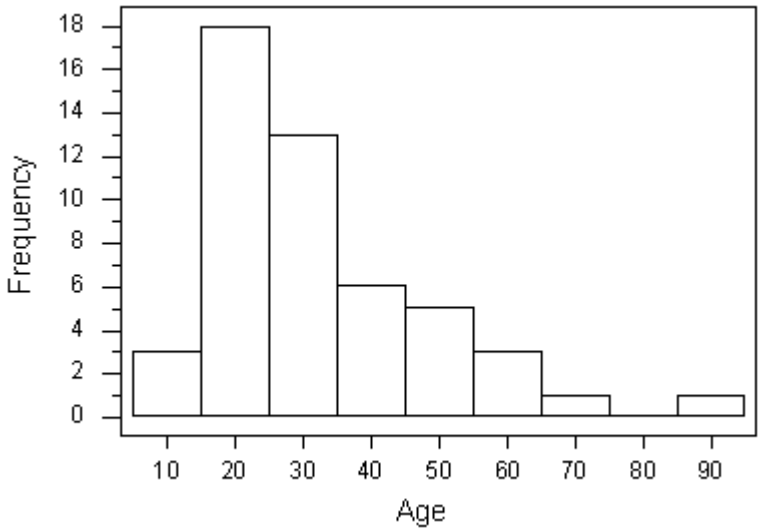
✗ This is not quite right. Remember that the ***range*** is the length from lowest value to highest value. So, for this distribution, the range is approximately 80. But remember that the ***median*** is the middle value. Half the people must be older than the median, so the median will be much less than 80, because all but one of the people were younger than 80. (B) is the correct answer.

E : 0

✗ This is not quite right. Remember that the ***range*** is the length from lowest value to highest value. So, for this distribution, the range is approximately 80. But remember that the ***median*** is the middle value. Half the people must be older than the median, so the median will be much less than 80, because all but one of the people were younger than 80. (B) is the correct answer.

Question (4)

Here again is the histogram showing the distribution of 50 ages at death due to trauma (accidents and homicides) that occurred in a certain hospital during a week.



A ***possible*** value of the median in this example is:

A: 23
B: 45
C: 50
D: 33
E: It is impossible to answer without seeing all of the data.

Feedback

A : 0

✗ This is not quite right. Remember that the ***median*** is the middle value. Only 21 people out of 50 could be younger than age 23, as represented by the total height of the two left bars on the display. So 23 is too low to be the median. (D) is the correct answer.

B : 0

✗ This is not quite right. Remember that the ***median*** is the middle value. Only 10 people out of 50 could be older than age 45, as represented by the total height of the four bars on the right side of the display. So 45 is too large to be the median. (D) is the correct answer.

C : 0

✗ This is not quite right. Remember that the ***median*** is the value of the middle person, not the middle of the range. Only 10 people out of 50 could be older than age 50, as represented by the total height of the four bars on the right side of the display. So 50 is too large to be the median. (D) is the correct answer.

D : 10

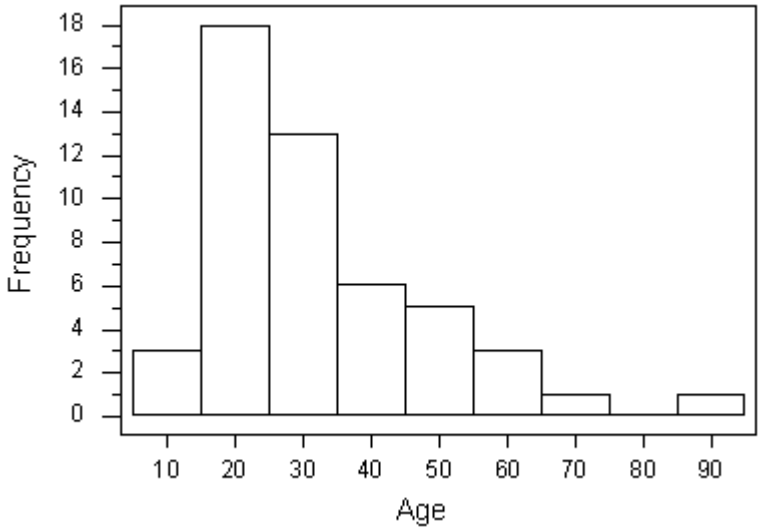
✓ Good job! The median is the ***middle*** value. There are fifty people in the data, so the median must be the age between the 25th youngest person and the 26th youngest person. These people will both be represented in the third bar from the left on the display. Since that bar represents ages from 25 to 35, the median could be any value between those ages.

E : 0

✗ This is not quite right. Although you cannot determine the exact median from the diagram, the question asked for a "possible" answer. Remember that the median is the ***middle*** value. There are fifty people in the data, so the median must be the age between the 25th youngest person and the 26th youngest person in the data. These people will both be represented in the third bar from the left on the display. Since that bar represents ages from 25 to 35, the median could be any value between those ages. (D) is the correct answer.

Question (5)

Here again is the histogram showing the distribution of 50 ages at death due to trauma (accidents and homicides) that were observed in a certain hospital during a week.



Assume that the largest observation in this dataset is 90. If this observation were wrongly recorded as 900, then:

A: Both the mean and the median will not change.
B: Both the mean and the median will change.
C: The mean will stay the same, but the median will change.
D: The mean will decrease, but the median won't change.
E: The mean will increase, but the median won't change.

Feedback

A : 0

✗ This is not quite right. It is true that the median won't change, because pulling 90 to 900 wouldn't change the 25th or the 26th value in the data. But remember that the ***mean*** tends to be "pulled" by an outlier. So moving the largest value, 90, farther to the right, to 900, would tend to "pull" the mean to the right, making it larger. (E) is the correct answer.

B : 0

✗ This is not quite right. It is true that the mean would change,

because mean tends to be "pulled" by outliers. But remember that the ***median*** is the "middle' value. Moving 90 to 900 wouldn't change the 25th or the 26th value in the data, so the median would be the same. (E) is the correct answer.

C : 0

✗ This is not quite right. Remember that the ***mean*** tends to be "pulled" by an outlier. So moving the largest value, 90, farther to the right, to 900, would tend to "pull" the mean to the right. So the mean would change. But the ***median*** is the "middle" value. Pulling 90 to 900 wouldn't change the 25th or the 26th value in the data, so the median would be the same. (E) is the correct answer.

D : 0

✗ This is not quite right. It is true that the median won't change, because moving 90 to 900 wouldn't change the 25th or the 26th value in the data. But remember that the ***mean*** tends to be "pulled" by an outlier. Moving the largest value, 90, farther to the right, to 900, would tend to "pull" the mean to the right, making it increase, not decrease. (E) is the correct answer.

E : 10

✓ Good job! The mean will increase, because the mean tends to be "pulled" by an outlier. So moving the largest value, 90, farther to the right, to 900, would tend to "pull" the mean to the right. The ***median*** wouldn't be affected, because moving 90 to 900 wouldn't change the 25th or the 26th value in the data.

Question (6)

The following data, on the number of children ever born per 1,000 women, is from the *Current Population Reports -Fertility of American Women: 2008*

Number of Children	None	One	Two	Three	Four	Five or More

Frequency	45.7%	17.0%	21.1%	10.6%	3.9%	1.8%

What is the modal number of children born to American women, in 2008?

A: Zero
B: One
C: Two
D: Three
E: Unable to Determine

Feedback

A : 10

✔ Correct! The most commonly occurring value is none; therefore, zero is the modal number of children born to American women, in 2008.

B : 0

✘ That's not quite correct. Remember that the mode is the most frequently occurring value; therefore, the correct answer is (A), since none is the most frequently occurring score.

C : 0

✘ That's not quite correct. Remember that the mode is the most frequently occurring value; therefore, the correct answer is (A), since none is the most frequently occurring score.

D : 0

✘ That's not quite correct. Remember that the mode is the most frequently occurring value; therefore, the correct answer is (A), since none is the most frequently occurring score.

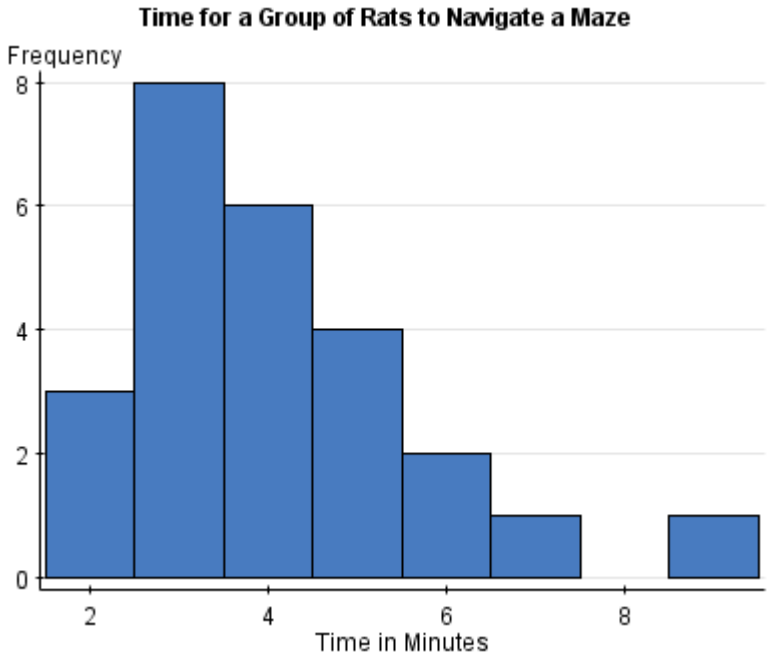
E : 0

✘ That's not quite correct. There is enough information to determine the mode. Remember that the mode is the most frequently occurring value; therefore, the correct answer is (A), since none is the most frequently occurring score.

The following 5 questions relate to the same histogram, shown below.

Question (7)

This histogram shows the times, in minutes, required for 25 rats in a animal behavior experiment to successfully navigate a maze.



What percentage of the rats navigated the maze in less than 5.5 minutes?

A: 34%
B: 60%
C: 68%
D: 70%
E: 84%

Feedback

A : 0

✗ That's not quite right. From the histogram we can find that 3 + 8 + 6 + 4 = 21 out of the 25 rats fall below 5.5 minutes (first four bars). Since the question asks about the percentage of

observations, the answer is 21/25 * 100 = 84%. (E) is the correct answer.

B : 0

That's not quite right. From the histogram we can find that 3 + 8 + 6 + 4 = 21 out of the 25 rats fall below 5.5 minutes (first four bars). Since the question asks about the percentage of observations, the answer is 21/25 * 100 = 84%. (E) is the correct answer.

C : 0

That's not quite right. From the histogram we can find that 3 + 8 + 6 + 4 = 21 out of the 25 rats fall below 5.5 minutes (first four bars). Since the question asks about the percentage of observations, the answer is 21/25 * 100 = 84%. (E) is the correct answer.

D : 0

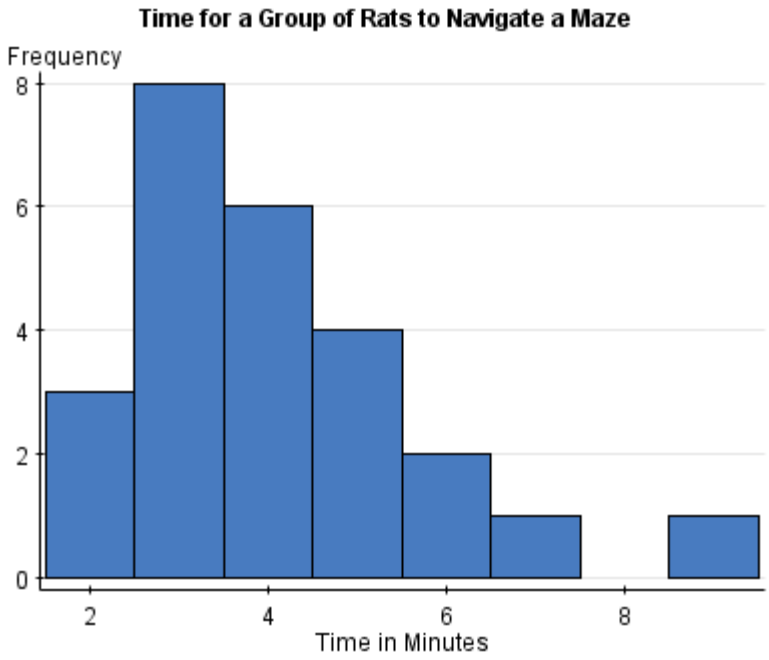
That's not quite right. From the histogram we can find that 3 + 8 + 6 + 4 = 21 out of the 25 rats fall below 5.5 minutes (first four bars). Since the question asks about the percentage of observations, the answer is 21/25 * 100 = 84%. (E) is the correct answer.

E : 10

Good job! From the histogram we can find that 3 + 8 + 6 + 4 = 21 out of the 25 rats fall below 5.5 minutes (first four bars). Since the question asks about the percentage of observations, the answer is 21/25 * 100 = 84%.

Question (8)

Here again is the histogram showing the times, in minutes, required for 25 rats in a animal behavior experiment to successfully navigate a maze.



Which of the following best describes the shape of the histogram?

A: Symmetric
B: Left-skewed with no outliers
C: Right-skewed with no outliers
D: Left-skewed with a possible outlier
E: Right-skewed with a possible outlier

Feedback

A : 0

✗ That's not quite right. Note that the histogram displays a tail to the right (i.e., it is skewed right) with one "isolated" observation at around 9 minutes, which is possibly an outlier. (E) is the correct answer.

B : 0

✗ That's not quite right. Note that the histogram displays a tail to the right (i.e., it is skewed right) with one "isolated" observation at around 9 minutes, which is possibly an outlier. (E) is the correct answer.

C : 0

✗ That's not quite right. Note that the histogram displays a tail to the right (i.e., it is skewed right) with one "isolated"

observation at around 9 minutes, which is possibly an outlier. (E) is the correct answer.

D : 0

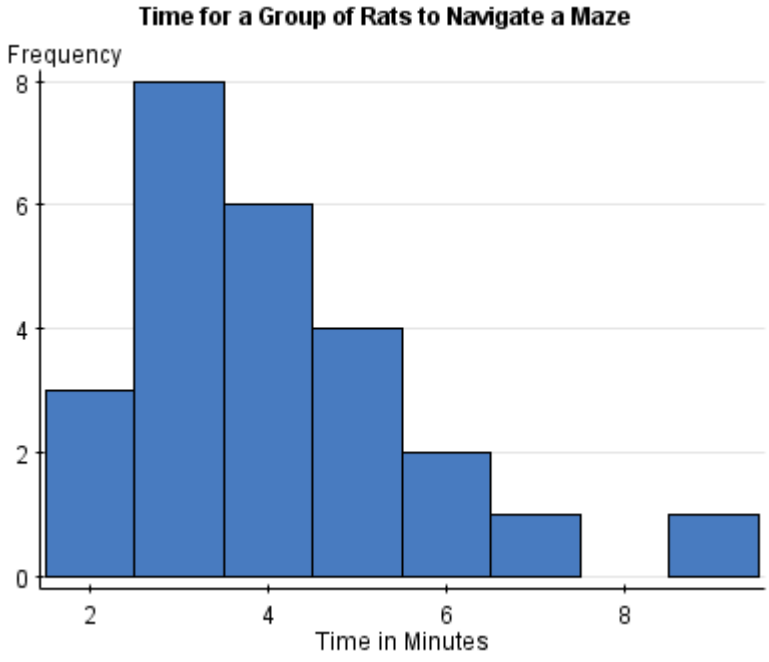
That's not quite right. Note that the histogram displays a tail to the right (i.e., it is skewed right) with one "isolated" observation at around 9 minutes which is possibly an outlier. (E) is the correct answer.

E : 10

Good job! The histogram displays a tail to the right (i.e., it is skewed right) with one "isolated" observation at around 9 minutes, which is possibly an outlier.

Question (9)

Here again is the histogram showing the times, in minutes, required for 25 rats in a animal behavior experiment to successfully navigate a maze.



For the data described by the above histogram,

A: the median will be smaller than the mean.
B: the median will be larger than the mean.
C: the median and the mean will be about the same.
D: the median and the range will be about the same.
E: both (A) and (D) are correct.

Feedback

A : 10

✔ Good job! When the distribution is skewed right, the mean is "pulled up" by the large observations in the "tail" of the distribution. The median, on the other hand, is only determined by the order of the observations, and is not affected by their magnitude. Therefore, in this case: mean > median.

B : 0

✘ That's not quite right. When the distribution is skewed right, the mean is "pulled up" by the large observations in the "tail" of the distribution. The median, on the other hand, is only determined by the order of the observations, and is not affected by their magnitude. Therefore, in this case: mean > median. (A) is the correct answer.

C : 0

✘ That's not quite right. When the distribution is skewed right, the mean is "pulled up" by the large observations in the "tail" of the distribution. The median, on the other hand, is only determined by the order of the observations, and is not affected by their magnitude. Therefore, in this case: mean > median. (A) is the correct answer.

D : 0

✘ That's not quite right. Remember that the ***range*** is the length from lowest value to highest value. So, for this distribution, the range is approximately 7. But remember that the ***median*** is the middle value. Half of the times must be greater than the median, so the median will be much less than 7, because all but two of the times were less than 7. (A) is the correct answer.

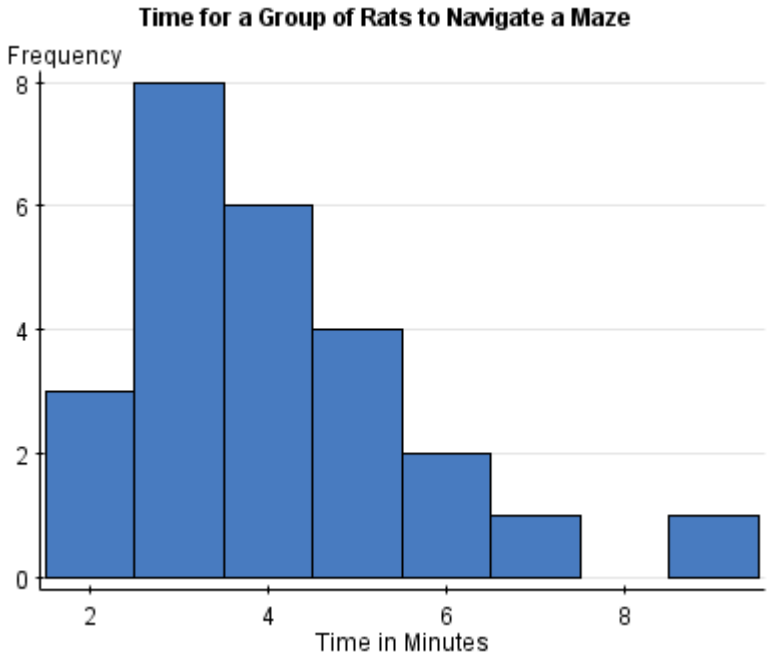
E : 0

✘ That's not quite right. Remember that the ***range*** is the length from lowest value to highest value. So, for this distribution, the

✗ range is approximately 7. But remember that the ***median*** is the middle value. Half of the times must be greater than the median, so the median will be much less than 7, because all but two of the times were less than 7. (A) is the correct answer.

Question (10)

Here again is the histogram showing the times, in minutes, required for 25 rats in a animal behavior experiment to successfully navigate a maze.



A ***possible*** value of the median in this example is:

A: 3.3
B: 3.9
C: 4.6
D: 5.5
E: It is impossible to answer without seeing all of the data.

Feedback

A : 0

✗ That's not quite right. There are 25 observations (odd number) and therefore the median is the 13th ranked observation. If we look at the histogram we can figure out that the 13th observation falls in the third interval (3.5– 4.5 minutes). The only option with such a number is 3.9 minutes. (B) is the correct answer.

B : 10

✓ Good job! There are 25 observations (odd number) and therefore the median is the 13th ranked observation. If we look at the histogram we can figure out that the 13th observation falls in the third interval (3.5– 4.5 minutes). The only option with such a number is 3.9 minutes.

C : 0

✗ That's not quite right. There are 25 observations (odd number) and therefore the median is the 13th ranked observation. If we look at the histogram we can figure out that the 13th observation falls in the third interval (3.5– 4.5 minutes). The only option with such a number is 3.9 minutes. (B) is the correct answer.

D : 0

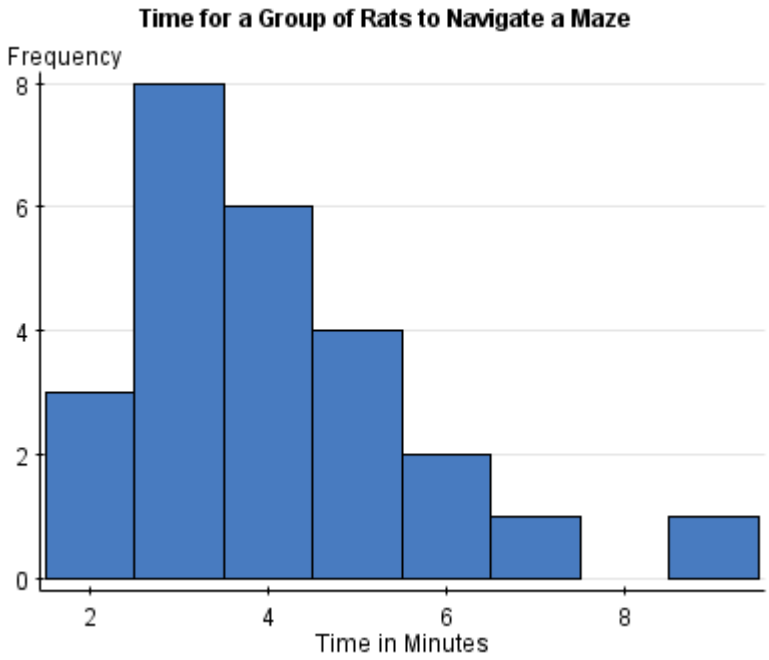
✗ That's not quite right. There are 25 observations (odd number) and therefore the median is the 13th ranked observation. If we look at the histogram we can figure out that the 13th observation falls in the third interval (3.5– 4.5 minutes). The only option with such a number is 3.9 minutes. (B) is the correct answer.

E : 0

✗ That's not quite right. There are 25 observations (odd number) and therefore the median is the 13th ranked observation. If we look at the histogram we can figure out that the 13th observation falls in the third interval (3.5– 4.5 minutes). The only option with such a number is 3.9 minutes. (B) is the correct answer.

Question (11)

Here again is the histogram showing the times, in minutes, required for 25 rats in a animal behavior experiment to successfully navigate a maze.



Assume that the largest observation in this dataset is 8.6 minutes. If this observation were wrongly recorded as 86, then:

A: Both the mean and the median will not change.
B: Both the mean and the median will change.
C: The mean will stay the same, but the median will change.
D: The mean will decrease, but the median won't change.
E: The mean will increase, but the median won't change.

Feedback

A : 0

✗ That's not quite right. The mean is affected by the actual values of the observations, while the median is determined by their order. Therefore, changing the largest observation from 8.6 to 86 will have an impact on the mean, while the median will stay the same.

✗ Actually, we can figure out the *exact* impact that the change will have on the mean: By changing 8.6 to 86, the sum of all the observations is increased by 86 - 8.6 = 77.4, and therefore the mean of the 25 observations is increased by 77.4 / 25 = 3.1 minutes. (E) is the correct answer.

B : 0

✗ That's not quite right. The mean is affected by the actual values of the observations, while the median is determined by their order. Therefore, changing the largest observation from 8.6 to 86 will have an impact on the mean, while the median will stay the same.

Actually, we can figure out the *exact* impact that the change will have on the mean: By changing 8.6 to 86, the sum of all the observations is increased by 86 - 8.6 = 77.4, and therefore the mean of the 25 observations is increased by 77.4 / 25 = 3.1 minutes. (E) is the correct answer.

C : 0

✗ That's not quite right. The mean is affected by the actual values of the observations, while the median is determined by their order. Therefore, changing the largest observation from 8.6 to 86 will have an impact on the mean, while the median will stay the same.

Actually, we can figure out the *exact* impact that the change will have on the mean: By changing 8.6 to 86, the sum of all the observations is increased by 86 - 8.6 = 77.4, and therefore the mean of the 25 observations is increased by 77.4 / 25 = 3.1 minutes. (E) is the correct answer.

D : 0

✗ That's not quite right. The mean is affected by the actual values of the observations, while the median is determined by their order. Therefore, changing the largest observation from 8.6 to 86 will have an impact on the mean, while the

median will stay the same.

Actually, we can figure out the *exact* impact that the change will have on the mean: By changing 8.6 to 86, the sum of all the observations is increased by 86 - 8.6 = 77.4, and therefore the mean of the 25 observations is increased by 77.4 / 25 = 3.1 minutes. (E) is the correct answer.

E : 10

Good job! The mean is affected by the actual values of the observations, while the median is determined by their order. Therefore, changing the largest observation from 8.6 to 86 will have an impact on the mean, while the median will stay the same.

Actually, we can figure out the *exact* impact that the change will have on the mean: By changing 8.6 to 86, the sum of all the observations is increased by 86 - 8.6 = 77.4, and therefore the mean of the 25 observations is increased by 77.4 / 25 = 3.1 minutes.

Question (12)

A survey was conducted to determine the number of servings of vegetables eaten, per day, by adolescents in the United States. The following table shows the results:

Number of Servings	Zero	One	Two	Three	More than 3
Frequency	5.7%	32.0%	34.0%	13.0%	15.3%

What is the modal number of vegetable servings eaten, per day, by adolescents in the United States?

A: Zero
B: One
C: Two
D: Three
E: More than 3

Feedback

A : 0

✗ That's not quite correct. Remember that the mode is the most frequently occurring value; therefore, the correct answer is (C), since two is the most frequently occurring score.

B : 0

✗ That's not quite correct. Remember that the mode is the most frequently occurring value; therefore, the correct answer is (C), since two is the most frequently occurring score.

C : 10

✓ Correct! The most commonly occurring value is two; therefore, two is the modal number of vegetable servings consumed, per day, by U.S. teenagers.

D : 0

✗ That's not quite correct. Remember that the mode is the most frequently occurring value; therefore, the correct answer is (C), since two is the most frequently occurring score.

E : 0

✗ That's not quite correct. Remember that the mode is the most frequently occurring value; therefore, the correct answer is (C), since two is the most frequently occurring score.

Please answer the question below. Your response will not be graded, but will be

available for your instructor to read.

What determines which numerical measures of center and spread are appropriate for describing a given distribution of a quantitative variable? Which measures will you use in each case?

Feedback

*** : 0**

✗ Thank you for your response. It has been recorded and will be available for your instructor to read.

✗ You did not enter a response for this question.